



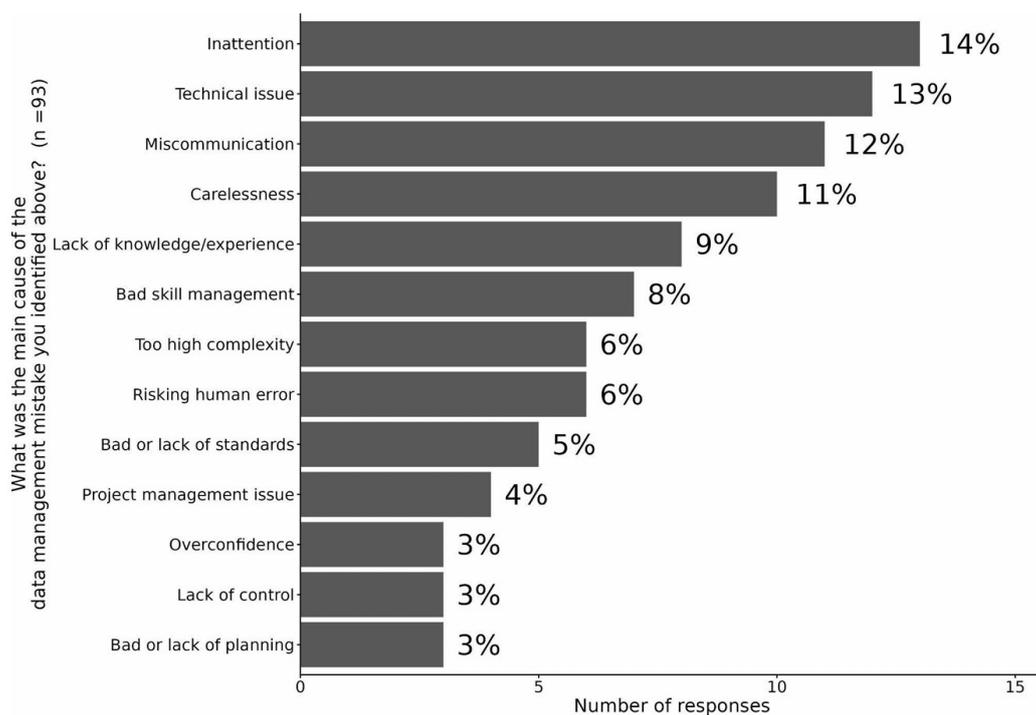
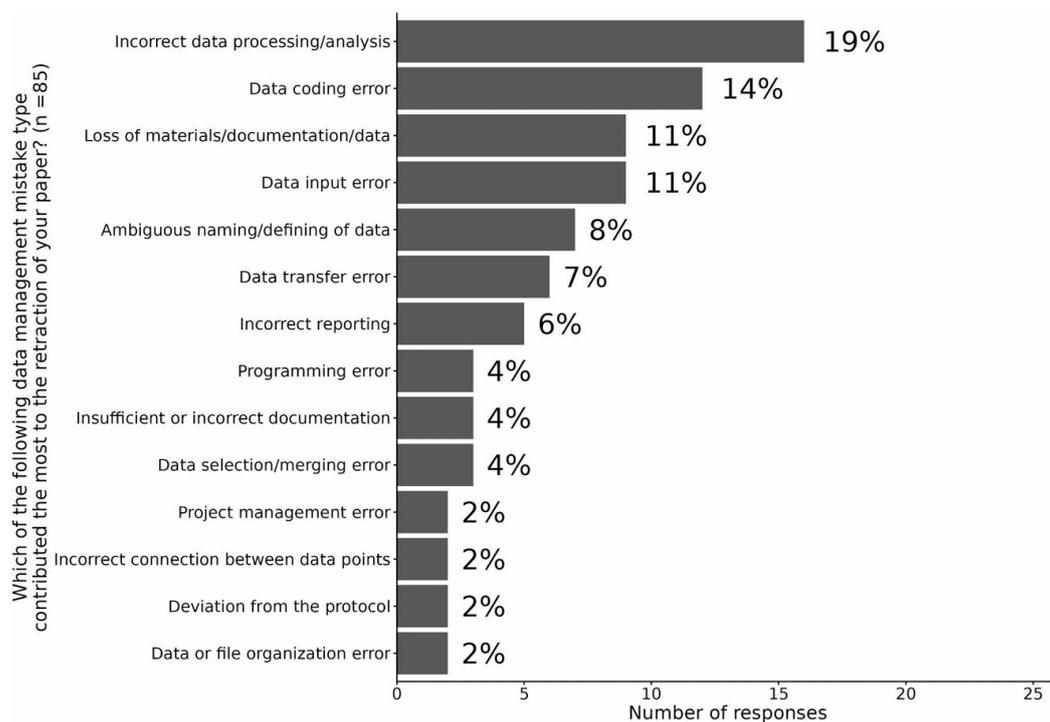
## 因为数据管理失误而被撤稿？应该如何避免？

论文被撤稿，除了因为学术不端，另一个重要的原因是非故意的失误 (honest error)。根据研究方法，以及筛选条件的不同，早期的研究显示，20%-60% 的撤稿是由非故意的失误造成的。如何帮助研究人员建立合理的制度降低这些失误的发生，成为了学术社区共同关心的问题。

在此之前，我们需要知道这些非故意的失误是怎么发生的。为此，来自于匈牙利和美国的联合小组通过向问卷的调查方式 [1]，向我们展示了哪些数据管理失误最容易导致论文被撤稿，以及这些失误是由于什么原因而产生的。

他们从 2018 年版本的 Retraction Watch Database 提取了 36773 篇被撤稿论文的信息，通过数据库的标签找出其中 5816 篇由于非故意的数据管理失误而被撤稿的文章。他们随后向论文的通讯作者发送邮件，邀请对方接受一项问卷调查，以询问具体发生了哪些数据管理失误，其原因，以及作者在事后的研究中进行了哪些改善措施。

246 名受访者回应了调查问卷，最后 97 人的问卷被纳入了分析。结果显示，数据处理/分析错误是数据管理错误中最常见的，约占 19%，其次是数据编码错误 (data coding error) (占 14%)，数据文件丢失 (占 11%)，数据输入错误 (占 11%)，以及数据命名不当 (ambiguous naming) (占 8%) 等。而造成这些失误的常见原因包括：不专心 (inattention) (占 14%)，技术性问题 (占 13%)，沟通失误 (占 12%)，粗心大意 (占 11%)，以及缺乏经验 (占 9%) 等。



对于如何避免这些失误的再次发生，受访者给出了不同的建议，包括落实数据处理者的责任，提供必要的培训，对数据进行更严格的验证核查等。此外，他们也希望期刊给出更明确的说明，或者指引，以明确哪些失误需要被撤稿，哪些可以通过订正给以修订。

撤稿对于研究人员来说，通常被认为是一件极为严峻的事情。47% 的受访者在回答“撤稿给他们带来多大程度的压力”时选择了最严重的选项（0-6 程度分级中选择了 6）。但我们需要承认，有些撤稿是因为非故意的失误所造成。希望这些前人的经验能帮助我们的研究人员更好的管理数据。

[1] [10.31222/osf.io/5t4xg](https://doi.org/10.31222/osf.io/5t4xg)

---

This article is licensed to the 5GH Foundation under a CC BY-NC-ND 4.0 International License